

Digital Searching to Digital Reading

Most searching is now online. Will most reading be online?

Online is popular and useful.

Technical obstacles are no longer the main problem.

Economic obstacles seem to be at least partially overcome.

What's left are the legal obstacles, and even those are being worked on, if not solved.

But how much will people want to read on line? And what will it mean to them?

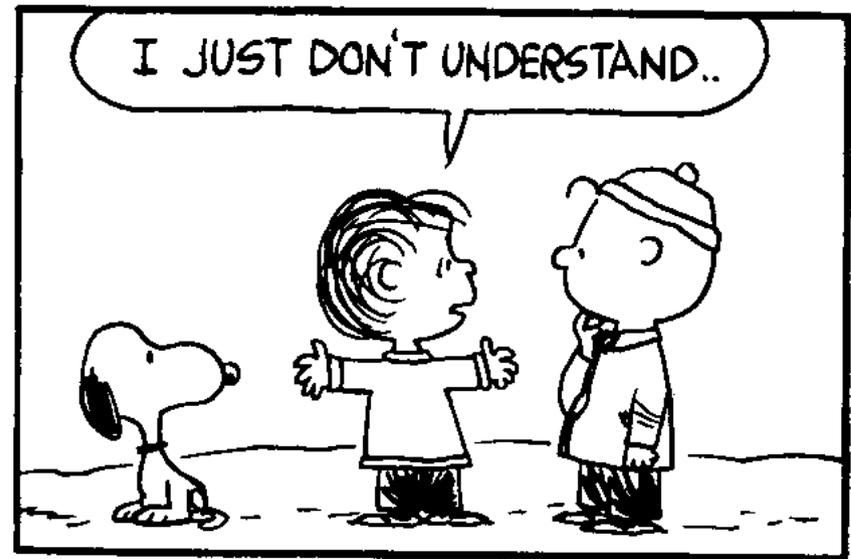
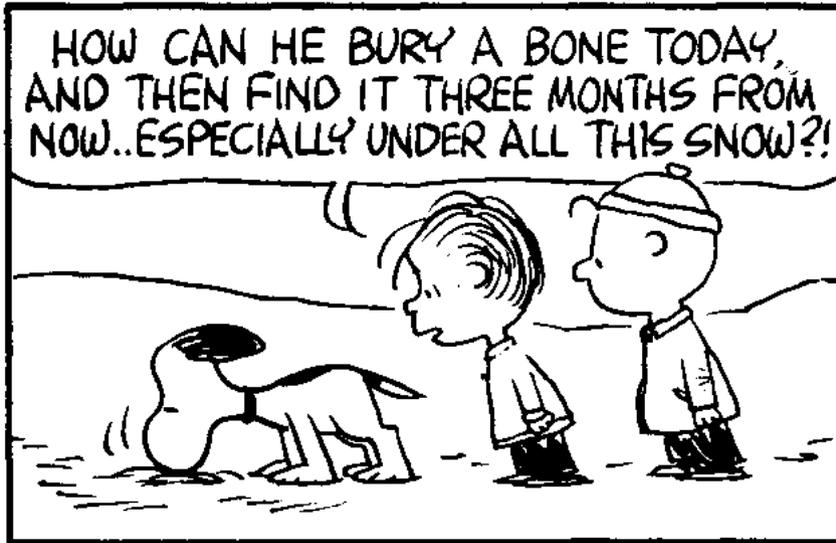
Everything Digital

“every child can stretch a hand across a keyboard and reach every book ever written, every painting ever painted, every symphony ever composed.” -- Bill Clinton’s State of the Union message, January 1998.

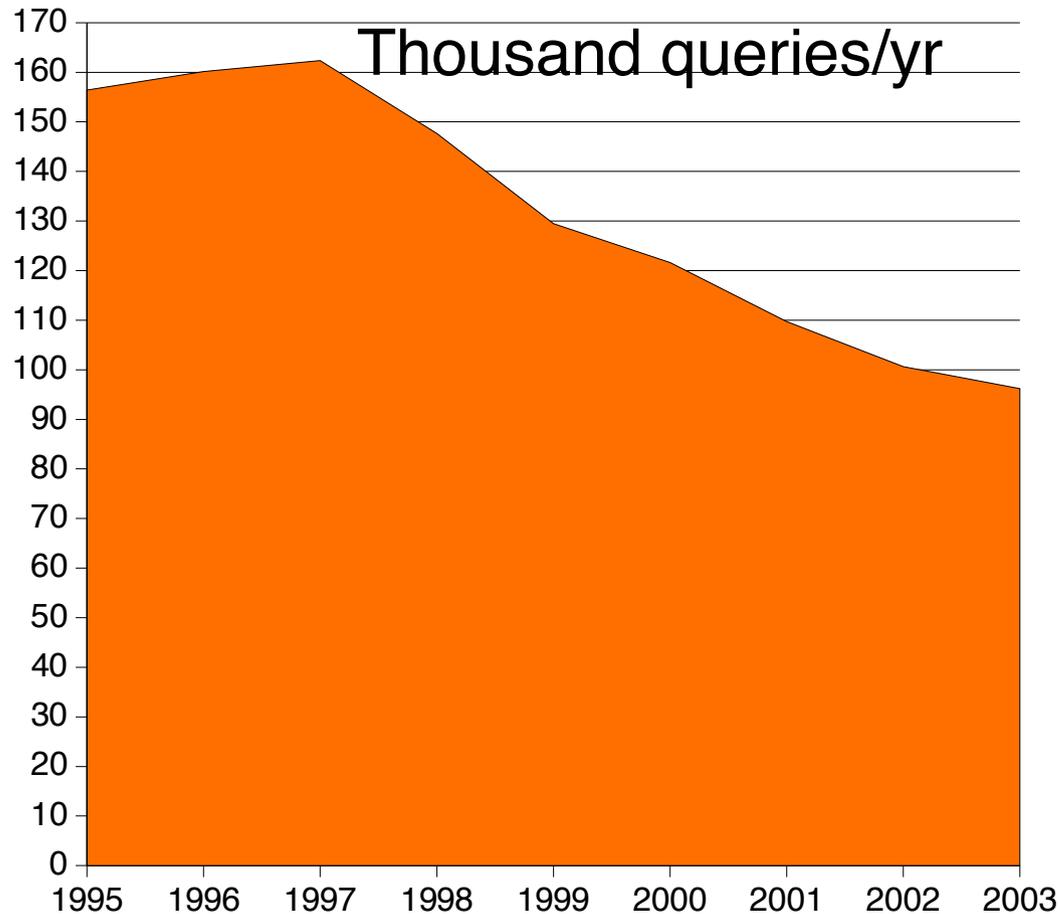
Similarly: H. G. Wells, *World Brain*, “There is no practical obstacle whatever now to the creation of an efficient index to *all* human knowledge, ideas and achievements, to the creation, that is, of a complete planetary memory for all mankind.”

“If the human race has produced since the invention of movable type a total record, in the form of magazines, newspapers, books, tracts, advertising blurbs, correspondence, having a volume corresponding to a billion books, the whole affair, assembled and compressed, could be lugged off in a moving van.” – Vannevar Bush, *As We May Think*.

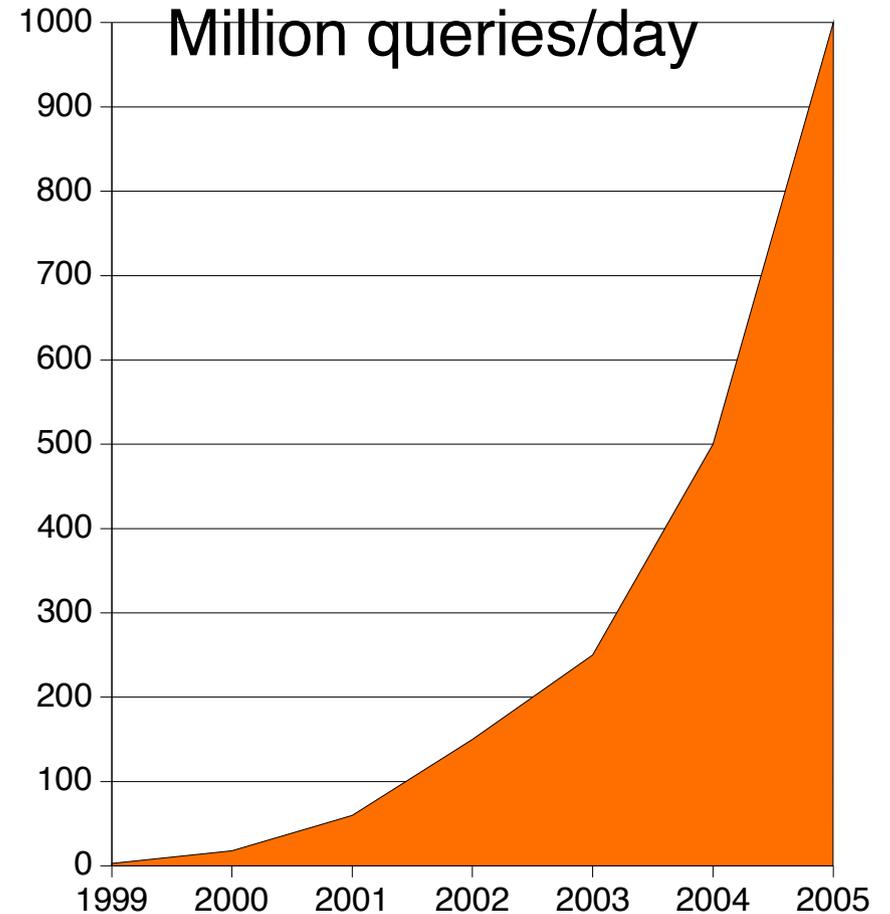
Life Before Google



Reference queries vs. Google



ARL average library, ca. 100,000 reference queries per year



Google about 1B queries/day;
2-4B all search engines.

For each query at an ARL reference desk, Google gets 10,000!!

Quantity, quality, and books

People complain that the web is full of junk; when Ian Irvine was head of Elsevier, he said that the web had the stuff rejected by his journals.

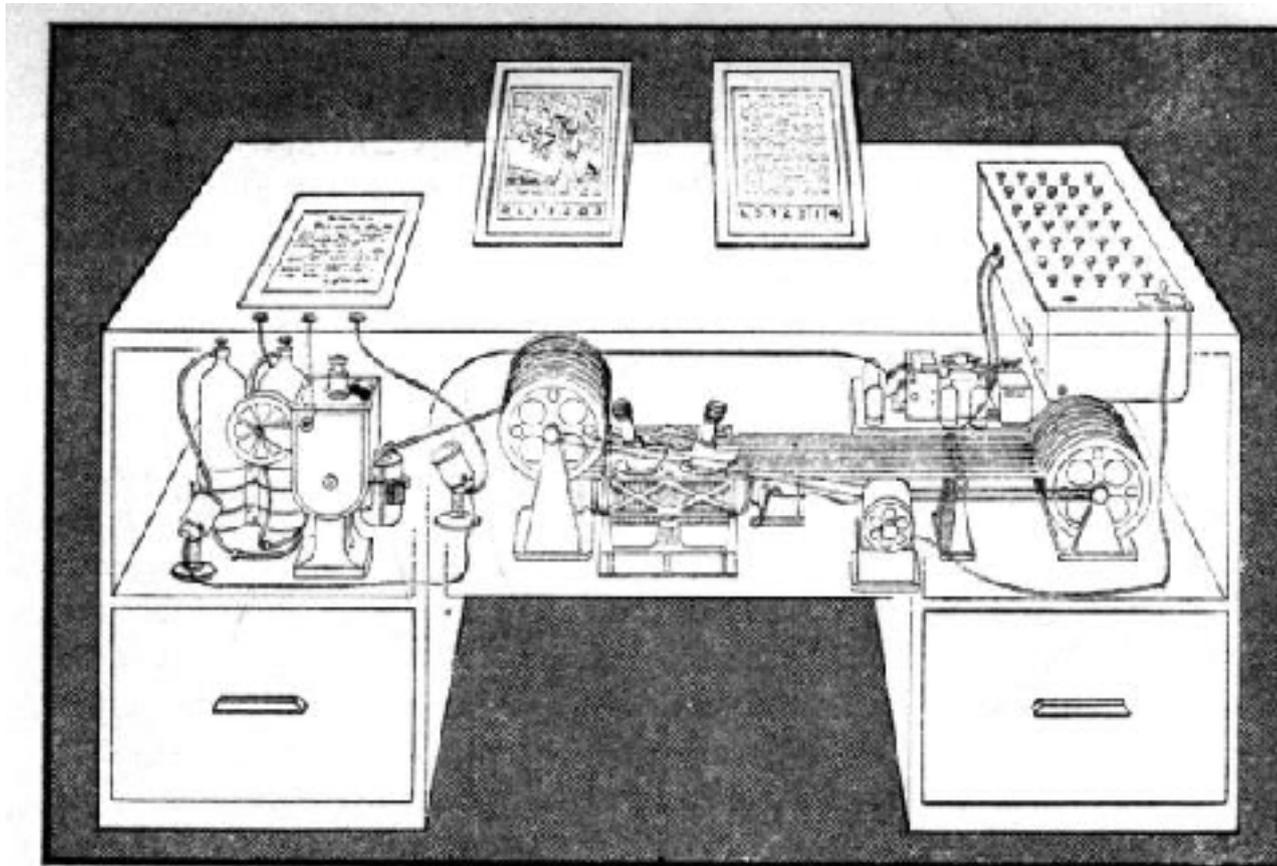
Is the web really that bad?

If we scan books, can we satisfy the critics?

How could we scan a lot of books?

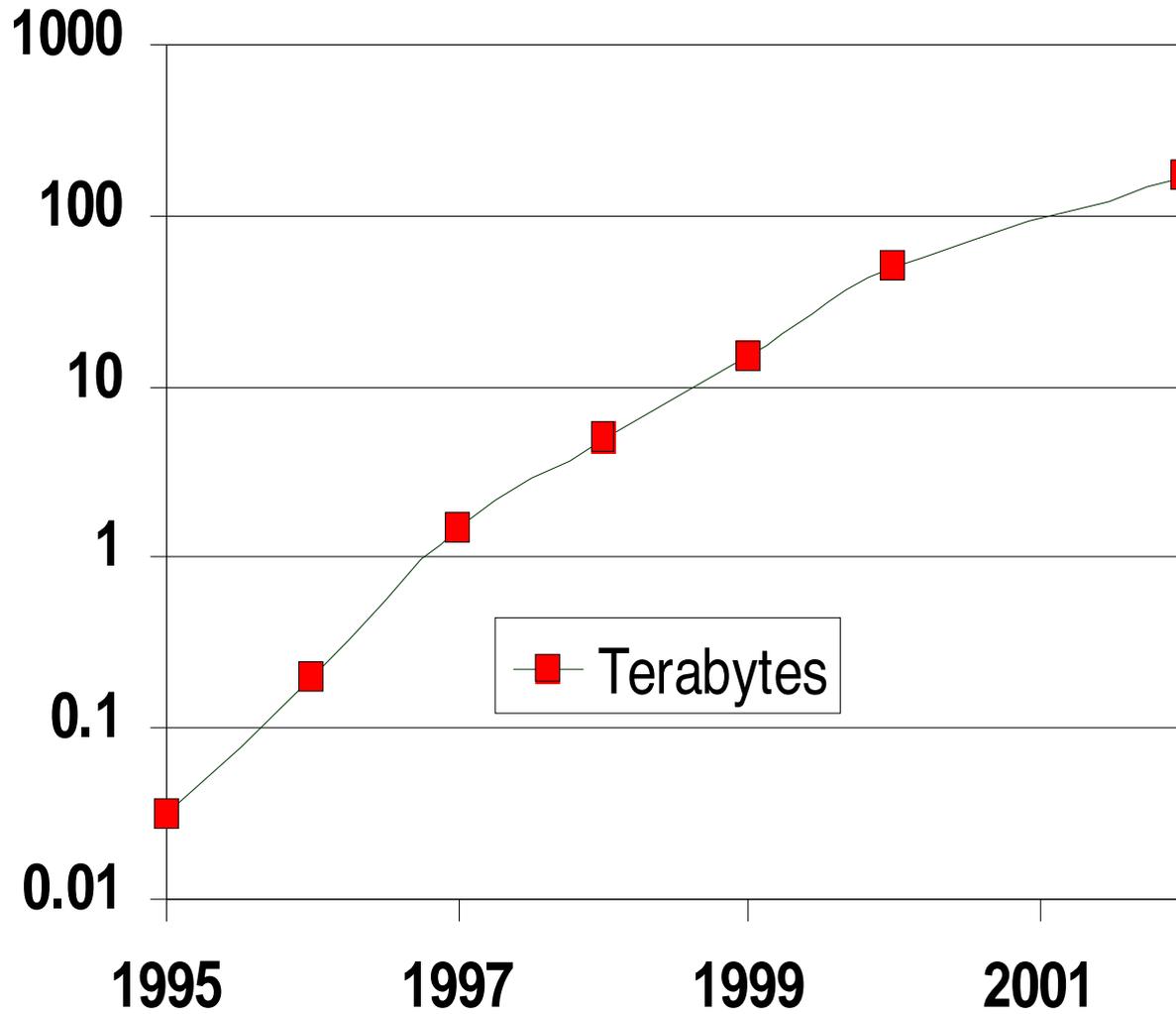
If we did have all these books online, how would we read them?

Bush's memex



As visualized by Life Magazine in 1945.

Size of the Web



Digital reading in libraries

Library of Congress, calendar year 2000 statistics:

	Items sent out	Approx. bytes
Website	900M	9 TB
Reading rooms	1.6M	1 TB

What to keep: lessons from history

Once upon a time libraries didn't give full respect to:

- Vernacular literature (before the Renaissance)

- Plays, instead of poetry

- Non-European languages

- Films and television scripts and recordings

Distinctions between libraries, archives and museums are eroding.

Undergraduates are using primary materials online, which they would not have been able to use on paper; even in schools some of these are useful.

As time goes on it is cheaper to collect but more expensive to select; it is cheaper to search and more expensive to organize.

Are digital libraries useful?

Despite a large investment of government money, few actual evaluations have been run; and people keep complaining about trash online.

A tentative result from Perseus suggested that students with lots of digital information wrote better bibliographies, but not better papers (Gary Marchionini).

Same conclusion in the 1980s at Stevens Inst. of Tech: computers help you search, but not think (Ed Friedman).

You would think, given the entire dependence of every high school and college student on Google, that somebody would have measured the result, but I don't know of such a study.

Google vs. ACM DL

Query: *neural nets*

ACM: 554 hits

Bounds for the computational
power & learning complexity..

Neural networks & open
texture

Efficient simulation of finite
automata....

Parallel construction of
minimal perfect hashing ...

Google: 131,000 hits

Lecture notes from Msc course
on neural nets

Neural networks at PNNL

Old neural net FAQ

FAQ for comp.ai.neural-nets

ACM dates 1991-1993, Google 1995-2001.

On balance Google pages better as an introduction; ACM hits too
specialized (ACM DL does not have monographs).

Google vs. ACM DL

Query: *rsa cryptography*

ACM: 12 hits

Hardware speedups in long integer multiplication.
Dynamically reconfigurable architecture for image proc.
Representation of ASN.1 in APL nested structures
Architectural tradeoff in implementing RSA procs.

Google: 117,000 hits

RSA Laboratories cryptography FAQ
RSA Labs algorithm simulation center (Javascript)
RSA Cryptography Today FAQ
RSA cryptography spec 2.0

Again, the ACM hits are very specialized; as an introduction the pages found by Google are better.

Google vs. Art Index

Query: *paleography*

Art Index: 72 hits

Cuneiform: The Evolution of a
Multimedia Cuneiform
Database

Une Priere de Vengane sur
une Tablette de Plomb a
Delos.

More help from Syria:
introducing Emar to biblical
study

The death of Niphururiya and
its aftermath

Google: 21,100 hits

Manuscripts, paleography,
codicology, introductory
bibliography

Ductus: an online course in
paleography

BYZANTIUM: Byzantine
Paleography

Texts, manuscripts and
paleography

Google vs. Art Index

Query: *Raphael fresco*

Art Index: 15 hits

Sappho, Apollo, Neophythagorean theory, and numine afflatur in Raphael's fresco of the Parnassus

Accidentally before, deliberately after (Raphael's School of Athens)

Raphael's Disputa: medieval theology seen through the eyes of Pico della Mirandola, and the possible inventor of the program, Tommaso Inghirami

Raphael's use of shading revealed (restoration of the Parnassus in the Stanza Della Segnatura almost completed)

Google: 8,950 hits

Raphael: The School of Athens

WebMuseum: Raphael: the nymph Galatea

OnArt Posterstore: Art Photography Music Film Posters

Raphael: Olga's Gallery

Google vs. Art Index

Query: *Inuit sculpture*

Art Index: 7 hits

Williamson, Karen. Terra "Cognita":
The Land We Know

Graburn, Nelson H. H. White
evaluation of the quality of Inuit
sculpture

Tunis, Roslyn. The art of research:
Nelson Graburn and the aesthetics
of Inuit sculpture at the Phoebe A.
Hearst Museum of Anthropology

Routledge, Marie. Carving an
identity: Inuit sculpture from the
permanent collection.

Google: 17,900 hits

Canadian Inuit Sculpture - Indian
and Northern Affairs Canada

Arctic Raven Gallery of Native Art:
Inuit Sculpture

North: Landscape of the Imagination
/ Le Nord: paysage imaginaire

Inuit art, Sculpture and prints from
Houston North Gallery

Google vs. index services

I tried four questions in computer science and six questions in art history, Google against the ACM digital library and the Art Index. In general:

- Google has more general resources

- Google sometimes gets distracted

It is hard to find a query that the professional resources do well but that Google can not do at all.

Are lots of books going online?

Project Gutenberg: more than 10,000 books, mostly English and American literature.

Gallica: 100,000 works of French literature and culture.

The Million Book Project: goal of a million books, covering many cultures but focussed on books in English, Chinese, and Indian languages)

Amazon 'search inside this book'

Google Print: planning to digitize all of the Michigan and Stanford libraries, and large collections from NYPL, Harvard, and Oxford

And commercial services such as Questia, ebrary, and Netlibrary, plus the many journals available from commercial publishers and nonprofits such as JSTOR, ACM, IEEE, High Wire, and so on.

The Million Book Project

Created by Raj Reddy of Carnegie-Mellon University; also led by Prof. N. Balakrishnan of the Indian Institutes of Sciences.

The US provides scanners, disks, and computers (about \$4.5M is committed); India provides labor (1 -2 thousand staff-years).

About 100 Minolta look-down scanners enable non-destructive black&white scanning of books at about one book per hour. With two shifts, for two years, this should scan 1 million books.

Scanning is 600 dpi, bitonal, with OCR and some image cleanup.

So far about 125,000 books have been scanned; 2/3 in India, 1/3 in China. About 1/3 in English, 1/3 in Chinese, 1/3 in various Indian languages.

Scanning

PLANT PROPAGATION

CHAPTER I

SEED AND SEEDLINGS

Around us, both at home and abroad, we observe that plants are propagated naturally by different means, but the majority are reproduced either by seeds or by spores. Of the latter we shall deal more fully when treating on



YOUNG FROND GROWING FROM PROTHALLUS. (Slightly Enlarged.)

p, Lower surface of prothallus; *rh*, root-hairs of prothallus;
f, young frond of Fern; *r*, root of Fern.

Ferns. Again, we observe in Nature that very wonderful and extraordinary means are employed to bring about cross-fertilisation and to prevent inbreeding, the progeny thus continuing to be reproduced without any weakening of their primary vigour.

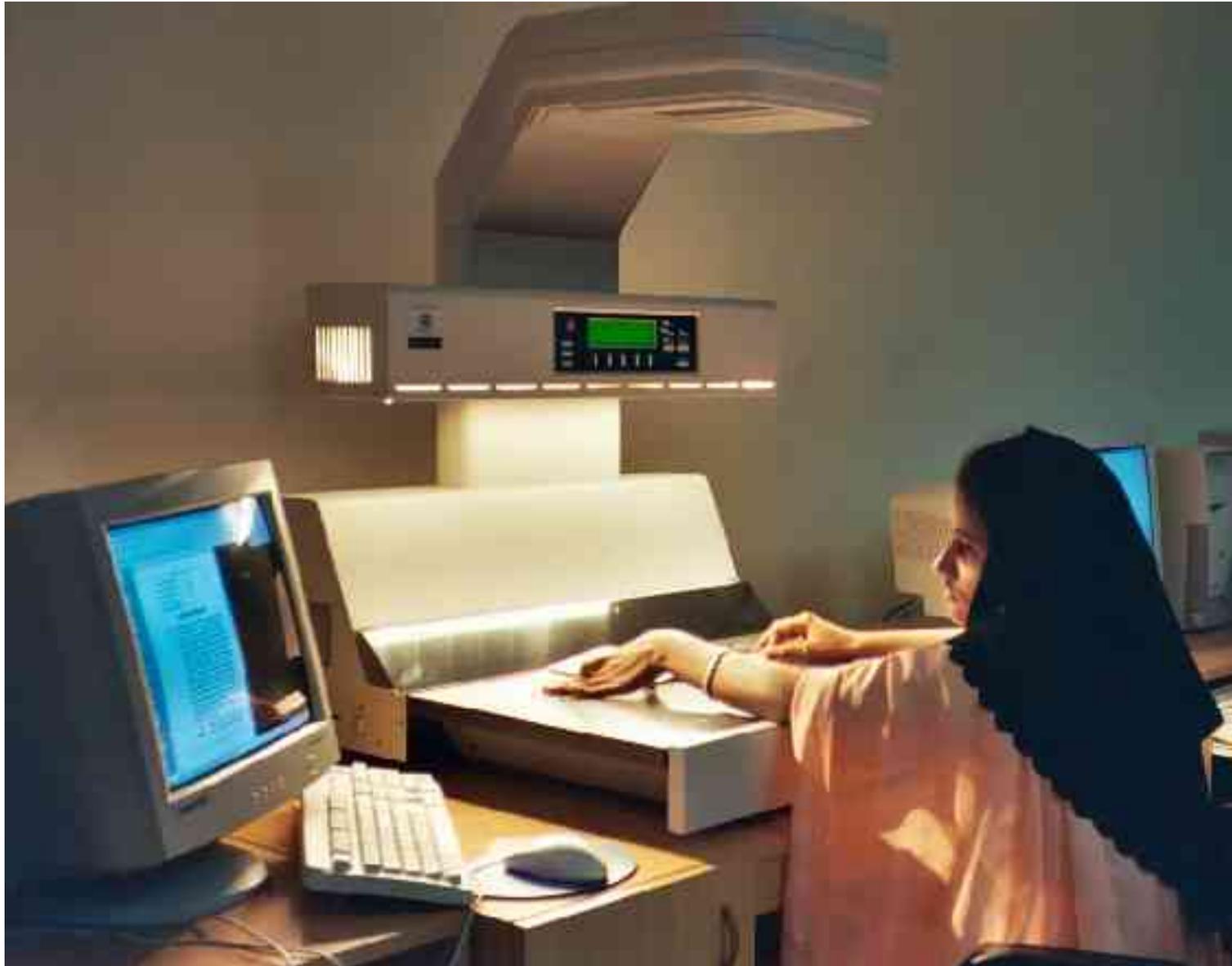


Scanning sites in India



Scanning sites in China

Minolta scanner



Rube Goldberg's scanner



This is a conventional flatbed scanner mounted upside down with a counterweight and foot pedal to allow the user to raise it and lower it. It's at the University of Georgia, and was entirely built of parts from scrapped library carts.

Robot page-turners



Digitizing Line

About \$300,000.



Kirtas Bookscan

About \$150,000.

Million Book equivalent cost

The scanners cost some \$7,000 and should be amortized over 10,000 books each, or under \$1 per book for hardware.

If the scanner does a book every two hours, with two people each paid 40 cents per hour, that's another \$1.60 or so.

So the actual cost is under \$3/book. We do better on best days; we also pay shipping but that's only 20 cents per book out.

Note: we can't afford selection.

Economics is not as bad a problem as it used to be.

Million Book Project Issues

Supply of books: people are afraid to lend (with reason).

Copyright law: most recent material not available

Access to results: poor bandwidth to India and China.

Coordination: not really adequate even within one country

Cataloging: need to train in OCLC

But nevertheless, we could scan about a million pages a day if we could get the books to scan.

Anyone who is renovating a library and putting a hundred thousand books in storage for a year – please talk to me!

Million Book Project progress

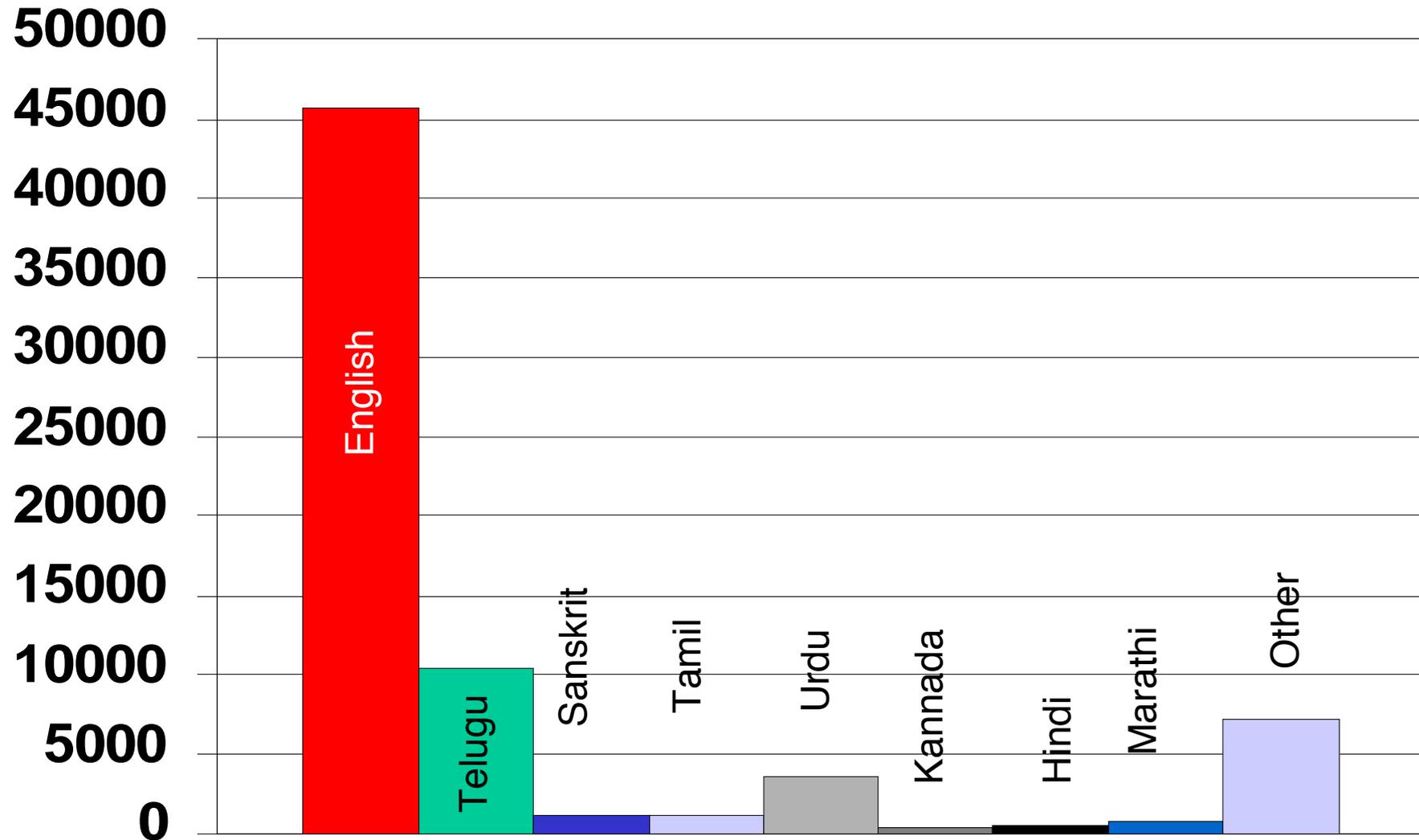
On a good day, a scanner does 5000 pages.

On an average day, 2000.

The difference is lack of books, poor management, and silly things like burned-out light bulbs. Evenly illuminated pages are essential for good quality scanning, but the light bulbs are special and have to be brought from Japan. By the way, they're so bright the operators wear sunglasses.

Large scanning operations (dozens of scanners) seem to be easier to get right than smaller ones; but politics mean we have small ones too.

Books by language (India)



Law: copyright, of course

The typical book in a research library is out of print, but not out of copyright. And, to quote AAUP: "The idea that ... anyone with a computer .. will be able to search the collections ... is enormously seductive. However, in our view it is built on a fundamental, broad-sweeping violation of the Copyright Act."

Administrative costs are the biggest problem: it's hard to find many copyright owners, hard to get permissions, and there's no simple economic way of paying the creators.

Compare, for example, mechanical recording rights rules or the way cable TV grew with protection from copyright claims.

Copyright renewals

In the United States,

- b) Books published before 1923 are public domain.
- c) Books published 1964-1988 are copyright for 95 years.
- d) Books published after 1988 are life of author + 70 yrs.

BUT

- d) Books published 1923-1963 had to be renewed 28 years later; otherwise they are public domain. Only perhaps 10% of the books that might have been renewed were renewed.

We had the index of renewals scanned and corrected.

Copyright renewals

You can access the records at

www.scils.rutgers.edu/~lesk/copyrenew.html

Warning: some interpretation needed. For example, the following as a result of looking for “Dickens Pickwick”:

RE-577-231

TITL: Pickwick papers.

AUTH: Charles Dickens, with an afterword by Steven Marcus CLNA: Steven Marcus (A) DREG: 28Jan92 ODAT: 17Dec64 OREG: A748735 OCLS: A LINM: NEW MATTER: afterword

Note that *only* the afterword is subject to copyright.

Counting

RLG lists some 36 million books, half in English.

About 2 million books published in the US before 1923.

About 1.9 million were published in the US 1923-1963, of which all but 5-8% were *not* renewed.

There are perhaps another 8M books in English in the RLG list from the UK and Canada; about 1.5M seem to be out of copyright (either author death date given in catalog as more than 70 years ago, or author born more than 140 years ago, or published more than 120 years ago).

So we might find 5 million books if we tried hard.

Many are not really serious books: leaflets, programs, ...

Million Book status

We're looking for:

- Libraries that are planning a major renovation which will mean that books have to go into storage for a year;
- Bookstores going out of business;
- Major weeding programs.

We're working on:

New technology for browsing and Indian language OCR/MT.

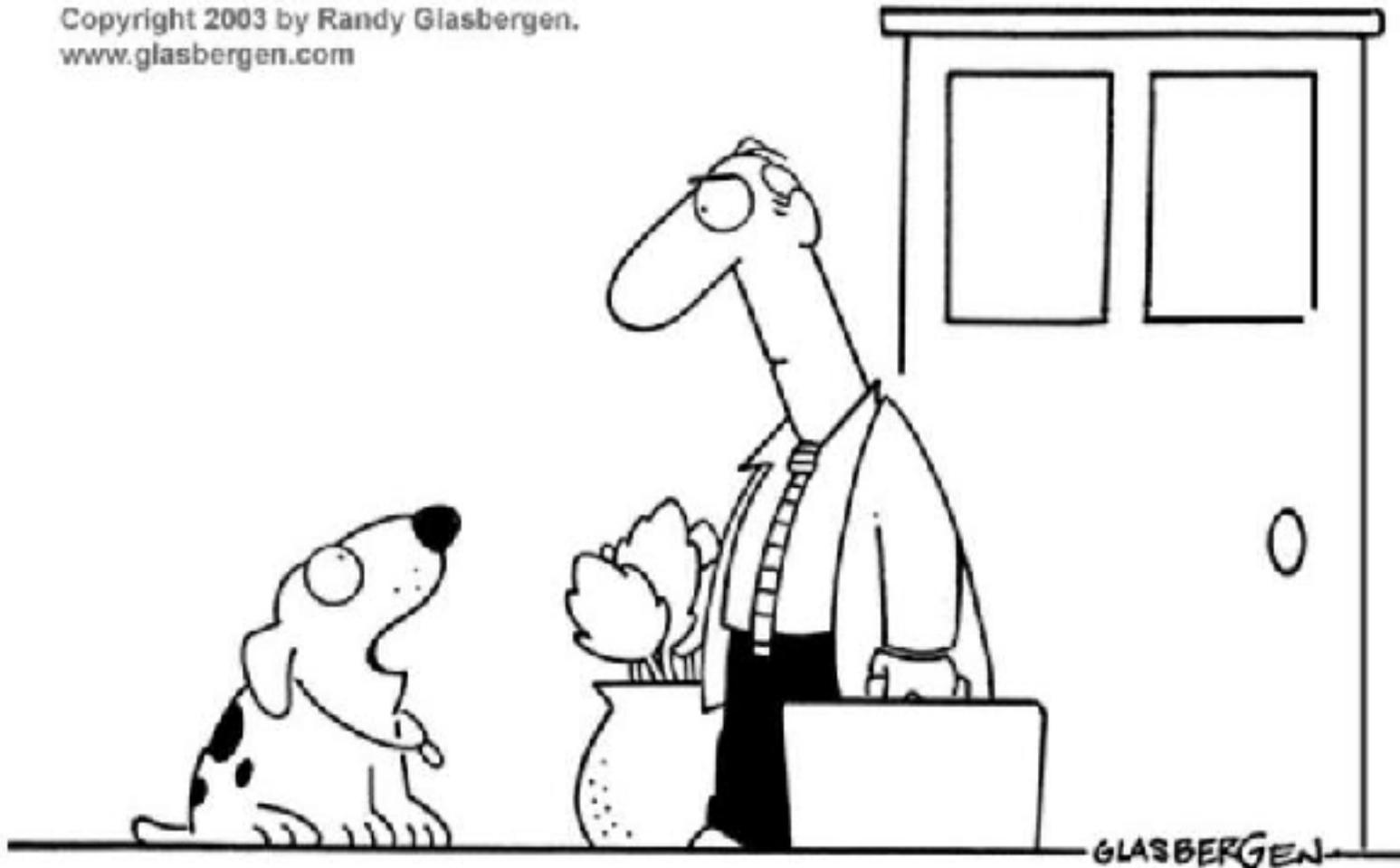
Better ways of moving books around, and scanning.

Writing for permissions.

Cooperating with those who hope to change the copyright law (Larry Lessig).

Biggest million book problem

Copyright 2003 by Randy Glasbergen.
www.glasbergen.com



Acceptance by librarians: part of the problem is delay, part a lack of local benefit, and part just wet feet. But not enough books.

“Before you can sell something, the customer has to like you! Did you remember to lick his face?”

Some think online reading is bad.

COMMENTARY Google and God's Mind
The problem is, information isn't knowledge.

(by Michael Gorman, president-elect of the American Library Association).

“The nub of the matter lies in the distinction between information (data, facts, images, quotes and brief texts that can be used out of context) and recorded knowledge (the cumulative exposition found in scholarly and literary texts and in popular nonfiction).

When it comes to information, a snippet from Page 142 might be useful. When it comes to recorded knowledge, a snippet from Page 142 must be understood in the light of pages 1 through 141 or the text was not worth writing and publishing in the first place.”

From the Los Angeles Times.... (December 17, 2004)

Better Book Browser (with Nina Wacholder)

Most web browsers are designed for typical web pages, which average 7K bytes.

We now have several sources of full-length books online, typically pre-1923 for copyright reasons.

Such books might be 500K-1M in length and have few illustrations.

At the moment, one typically has PDF. That lets you search and move by page number. Can one do better?

Why not PDF?

PDF allows search, and move by page; it does not help with an overview, unless the pages have images on them and thumbnails are useful.

It has a basic design principle of maintaining the visual image of each page, so it can't reformat to adapt to screen or window size.

Difficult to maintain orientation; is something you are looking for ahead of you or behind you?

If the book doesn't come with a summary or table of contents, nothing in PDF helps you.

Shneiderman & Marchionini 1988

"Today's electronic retrieval systems ... focus on coding, indexing, and cross-referencing (organization for retrieval) rather than on meaning, readability, and assimilation (organization for understanding)."

This was before the Web, but it was part of an argument for hypertext (specifically Hyperties).

Shneiderman has always argued for context, for systems that help people understand where they are in a task.

SuperBook – Egan and others, 1988

Quit	Help	Lookup	↑ Search ↓	Other
------	------	--------	------------	-------

<p style="text-align: center; background-color: black; color: white; margin: 0;">Table of Contents</p> <p>174 Journal of the American Chemical Society 68 Biochemistry *Pharmacology *Toxicology 10 *General Biochemistry 22 Enzymes 2 *Separation-purification-general characterization *Substrates-cofactors-inhibitors-water-coenzymes-products 7 Kinetics-mechanism-enzyme and coenzyme models 1 Co(III) Complex Promoted Phosphodiester Hydrolysis Location: Dep. Chem.; McGill Univ.; Montreal Text Experimental Section 1 Results and Discussion Conclusions</p> <hr/> <p>174 occurrences of "zinc"</p>	<p style="text-align: center; background-color: black; color: white; margin: 0;">Page of Text</p> <p style="text-align: right; background-color: black; color: white; margin: 0;">close</p> <p>Biochemistry Enzymes Kinetics-mechanism-enzyme and coenzyme models</p> <hr/> <p>Co(III) complexes can be freed hydrolytically. * Cobalt Complex Promoted Hydrolysis of BNPP. Our experimental data on $\text{cis-[Co(en)}_2\text{(OH)(OH}_2\text{)]}^{2+}$-promoted hydrolysis of BNPP can be explained in terms of the mechanism shown in Scheme 1. ##sch# We can rule out direct intermolecular Co(III)-hydroxide attack on BNPP, since the substitutionally inert cobalt complex $[\text{Co(en)}_2\text{(NE}_2\text{)(OH)}]^{2+}$ has minimal activity (Table 1). Hydrolysis of 2 must be more rapid than the rate of its formation, since under our experimental conditions, 2 equiv of p-nitrophenol is produced with no accumulation of the phosphononester intermediate 2. Sargason et al.⁹ showed by isotope-labeling experiments that the hydrolysis of the phosphononester bond in $[\text{Co(en)}_2\text{(OH)(OP(O)}_2\text{OC}_2\text{H}_4\text{NO}_2)]$ occurs by a four-membered ring formation analogous to the case in Scheme 1. Anation of $[\text{Co(en)}_2\text{(OH)(OH}_2\text{)]}^{2+}$ with phosphates has been studied by Lincoln.²¹ The k_2 step in Scheme 1 is the rate-determining step, since anation of the cobalt complex is faster than our observed hydrolysis reaction. $\text{cis-[Co(en)}_2\text{(OH)(OH}_2\text{)]}^{2+}$ is in rapid equilibrium with $\text{trans-[Co(en)}_2\text{(OH)(OH}_2\text{)]}^{2+}$ during the course of the phosphonester hydrolysis reaction. At neutral pH the</p>
--	--

1. zinc (174)

Found: zinc

Make alias Unalias

The Story of the Nations.

MODERN ENGLAND



This is our test book: Justin McCarthy's *Modern England*. A more accurate title would be something like "Political history of England, 1815-1835" or even "The first Reform Bill and the Abolition of the Slave Trade."

(The image is a PDF viewer)

IX.

THE GREAT REFORM BILL 219-253

Popular excitement—Brougham and Russell—Grey and Brougham—Napoleon at Elba—The expected measure—Lord John Russell's speech—Rotten boroughs—The Government proposal—The New Franchise—Sir Robert Inglis—Sir Robert Peel—Peel against the Bill—Daniel O'Connell's speech—The policy of the Tories—The second reading—Parliament dissolves—The new Parliament.

X.

THE REFORM BILL AGAIN 254-284

Parliamentary obstruction—A Committee of obstruction—The House of Lords—Defeat in the Lords—Serious riots—If the King holds out—Execution of rioters—"The Waverers"—Lord Grey's resignation—The King's Unpopularity—The Royal Assent—Some great Conservatives—"Finality"—Farewell to Lord Grey.

An old-style table of contents, with phrases for each section.

which kept him a prisoner in his room. Two days after Huskisson had left him, Canning removed to the Duke of Devonshire's villa at Chiswick; and a few days after his life came to an end there. He died on the 8th of August, 1827, in the very room where Charles James Fox had died not so many years before. The whole nation mourned his untimely death. Canning was but little more than fifty-seven years old when his great career came to a close.

The Duke of Wellington was then called upon to form an Administration. Some measures of political importance marked the course of the Duke's tenure of office, but for the present we are chiefly concerned with one event which might almost be said to belong to the career of Mr. Canning. There were differences of opinion in the Cabinet on many questions; and especially on those with which the name of Canning was particularly associated. Huskisson made a speech at Liverpool in addressing his constituents there, which led to a serious dispute between Wellington and himself. Huskisson told, or was reported to have told, his constituents at Liverpool that he never would have taken office under the Duke of Wellington if he had not obtained from the Duke an ample guarantee that the policy of Canning on financial and other questions was to be faithfully carried out. Wellington, with his usual bluntness, repudiated any such idea; threw contempt upon the suggestion that any gentleman to whom he had tendered office could think of insisting on any such guarantee, or making a guarantee of any kind a condition of accepting a place in the new Administration.

What is browsing?

Dealing with an unarticulated problem?
Finding material not in response to an "information need?"
Basically, relying on recognition rather than recall.

"How did William Huskisson die?" is not browsing.
"Find an interesting book" doesn't sound like browsing, but at the moment we have no other way to do it.

Browsing is also a way of trying to select items for more detailed examination: there ought to be a tradeoff between time spent choosing what to read and time spent reading (in days of yore Don King reported 20% on choosing).

State of the art?

Marti Hearst: Tilebars
David Harper: Clusterbook

In both cases, a vector of boxes shows the linear position of word hits.

Attempts to visualize the results of searches; less valuable if you don't have a search term.



Figure 4. TileBar search on (patient medicine medical AND test scan cure diagnosis AND software program) with some distribution constraints.

I. ARMS AND THE MAN . PAGE 1-9.

Napoleon, 'The Iron Duke.'

II. ENGLAND'S "BENEVOLENT" DESPOT
10-35.

The Weyinger Pitt, Charles James Fox,
The Disappointments of Peace, Religious
Disabilities, The Democracy, The King
puts his foot down, The King's good
points, Death of George III.,

III. IN THE WAKE OF THE PEACE 36-65.

The Congress of Vienna, Division of the
spoils, The Holy Alliance, The Allied
Sovereigns, George Canning, The Prince
Regent, Needed Reforms, Lord Castlereagh,
Lord Eldon, Lord Castlereagh,
Informers.

IV. GEORGE IV: 66-87.

What was thought of the King, Queen
Caroline, The Duke of York, Peterloo,
The Public Meeting, The Meeting broken
up, "Some one had blundered", Lord
Eldon's Law, the right of Meeting.,

V. THE CATO STREET CONSPIRACY .

88-107.

The Cato Street Conspiracy,
Thistlewood, The Garret in Cato Street,
The Battle of Bonnymuir, Execution of
Thistlewood, "The world is not try

IN this volume and in the one to follow my purpose is to give an account of the social and political development of England since the opening of the century.

I do not attempt anything like a minute and detailed history of the events that followed each other during that time; and indeed my intention is rather to give something like a picture than to give to my readers a chronicle and a record.

I have endeavoured to describe each remarkable political and social development, and to group the statesmen and philanthropists of every order by whom each development was assisted in its progress.

I have, while keeping in view the order of historical succession, endeavoured to make the story of each great reform, political or social, a story complete in itself, and disentangled as far as possible from the cluster and confusion of events that were passing all around it, and exterior to it.

The true history of England during that long period of marvellous growth will be found to be the history of the country's progress in education, in science, and in the conditions that tend to make life useful, healthful, and happy.

Successive changes in administration, the rivalries of statesmen at home and abroad, the barren wars which spring from the competing ambitions of dynasties are, after all, but the accidental difficulties in the way of man's improvement; and while they cannot be denied, their proportionate representation in history are sometimes dealt with as if they were the main events of a story, and were entitled to occupy the largest space and foremost place in the picture drawn by the historian.

My purpose has been something different from this; my desire has been to describe the marvellous changes wrought by science and literature, by statesmen and philanthropists in the course of England during the wonderful century. It is now drawing to a close.

My wish has been to make my readers acquainted with the men who helped to bring about those changes, as well as with the nature, extent, and influence of the changes themselves.

Mr. Huskisson was one of those who attended. He had been paying a visit to his constituents in Liverpool, and, although in very feeble health, he had made up his mind to be present on the memorable occasion of the opening of this first completed railway. Before the train left Liverpool the railway authorities requested the company to keep their places in the carriages until the train reached its destination and a printed handbill, setting forth the request, was passed along among the travellers. It seems almost unnecessary to say that the request and caution were of no avail. The train stopped at a wayside station a few miles down the line, and several of the company immediately got out and indulged their curiosity by walking up and down and inspecting the outside of the carriages. Unluckily a friend of the Duke of Wellington and of Huskisson was seized with the idea that this would be a propitious moment to bring the two men together and get them to shake hands. The idea spread abroad, and both the Duke and Huskisson were quite willing to take this opportunity of renewing

their former friendship. The Duke advanced along the platform to meet Mr. Huskisson, who was approaching him, and held out his hand in cordial greeting. Before Huskisson had well time to take the proffered hand some alarm was caused by the reported approach of a locomotive, and a cry was raised to those who were standing outside admonishing them to get into the carriages again. Huskisson was standing by the open door of one of the carriages and was not quick in getting in, probably because of his physical weakness. The open door at which he stood was struck by the locomotive, and Huskisson received injuries so severe that he died almost immediately after being removed to a neighbouring parsonage. The fatal event, of course, cast the deepest gloom over the whole party, the Ministers were only induced not to break up the ceremonial at once by the fear that some terrible alarm might be spread over Manchester. This was again a season of alarm, and no one could tell what exaggerated fern bewildered rumour might not take if none of the members were to arrive at the place of their destination. Nothing seemed more probable than that an affrighted tale might be spread about the assassination of the Ministers of the Crown in a body, and might lead to the vilest disturbances. In any case it was earnestly urged upon the Ministers

An interface with a scrollbar, search box, and simple text display with search terms highlighted.

Huskisson

search

quit

Context-sensitive browsing

One advantage computers should have is topic-sensitive presentation; imagine, for example, a person reading the newspaper and looking for material relevant to economics, or politics, or entertainment, or whatever. Similarly, a reader of a book on English history might have a world view in which "Oxford" is the site of a university, the site of a car factory, or the site of a battle (or even a kind of cloth).

Today, we don't do this: each person browsing sees the same result.

Trying for phrases instead of words...

holy alliance, holy alliance
john ainslie, john ainslie, george iv
old carter, old carter
duke of wellington, policy of england
lord elton, lord john russell
state church, date of supremacy
lord john russell, sir robert inglis
reform bill, reform bill
lord ashley, public opinion

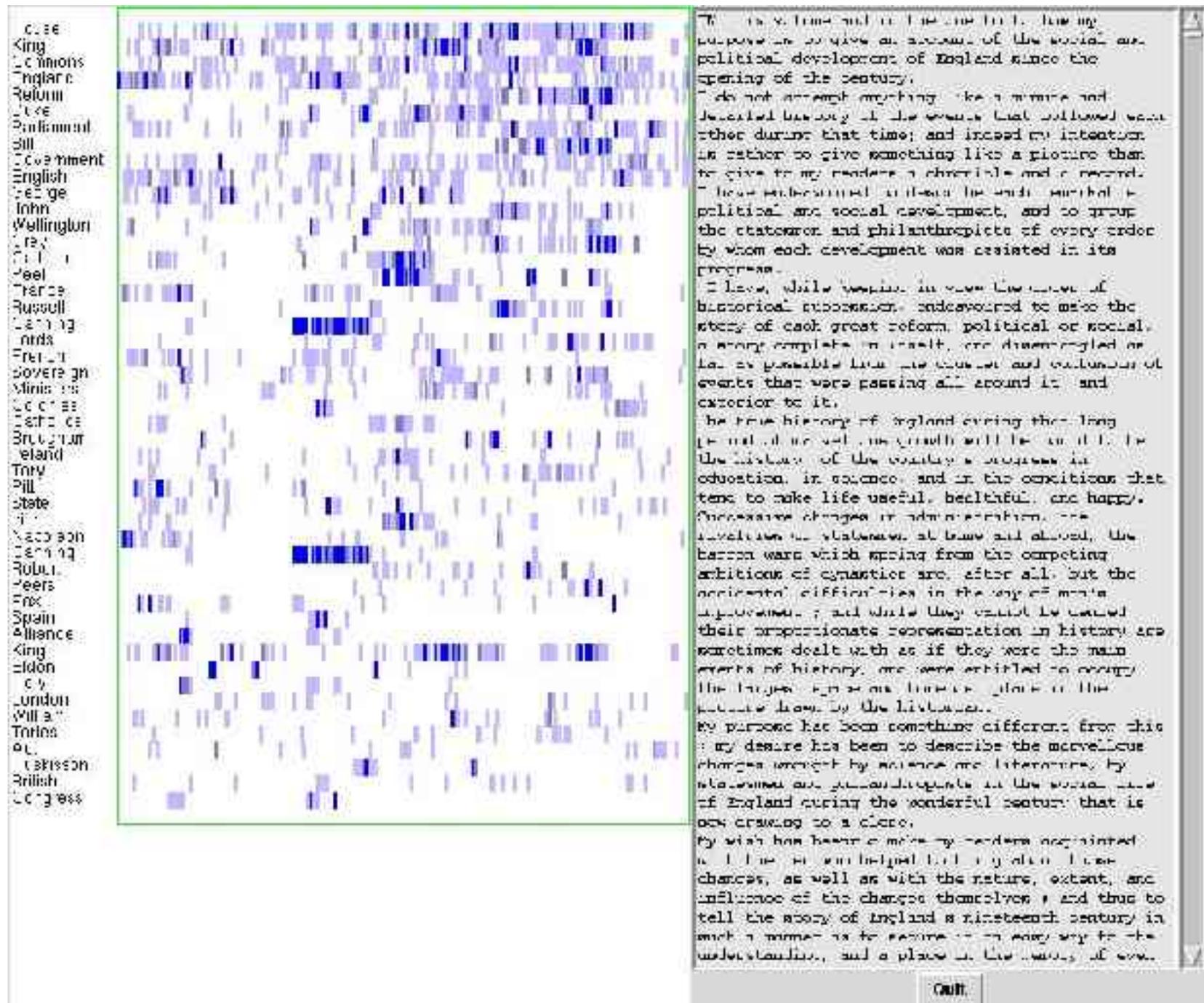
nineteenth century, {}
napoleon bonaparte, {}
turn and reverse, {} nineteenth century
battle of waterloo, {}
benevolent despot, condition of england
intellect of elton, condition of england
reform cause, william pitt
peace to europe, {}
policy of lord, {}
box and patt, {}
george iii, henry erskine
benevolent despot, royal intellect
george iii, george iii
benevolent despot, {}
treaty of vienna, madman than shakespeare
name of parliament, {}
prussian government, part catholic
holy alliance, russia from st.
reign of george, {}
holy alliance, duke of wellington
holy alliance, policy of england
lord elton, {}
prince regent, {}
lord brigiam, name of lord
popular reform, family name
name of lord, {}
lord eldon, {}
lord liverpool, home secretary
home system, {}
george iii, government for ireland
prince regent, {}
public opinion, {}
duke of york, {}
crater hunt, ready morning
universal suffrage, {}
hunt - heron, {}
chief constable, {}
chief constable, {}
name of parliament, {}
lord eldon, indictment against hunt
peterloo meeting, john bright
name of lord, {}
home office, home secretary
lord sidmouth, {}
date of lord, {}
manchester meeting, {}
battle of boyne, {}
reform bill, battle of boyne
name of lord, {}
house of commons, {}

IN this volume and in the one to follow my
purpose is to give a picture of the social and
political development of England since the opening
of the century. I do not attempt anything like a
minute and detailed history of the events that
followed each other during that time; and indeed
my intention is rather to give something like a
picture than to give a history of the events and
a record. I have endeavoured to describe each re-
markable political and social development, and to
group the statesmen and philanthropists of every
order by whom each development was assisted in
its progress.

I have, while keeping in view the rules of his-
torical succession, endeavoured to make the story
of each great reform, political or social, a story
in itself, - a story which would be as interest-
ing to the student and student of events that were
passing all around it, and superior to it. The true
history of England during that long period of
marvellous growth will be found to be the history

of the country's progress in education, in science,
and in the arts, and that that progress is a story
healthful and happy. Successive changes in ad-
ministration, the rivalries of statesmen at home
and abroad, the changes which are the result of
the competing ambitions of dynasties are, after all,
but the accidental difficulties in the way of man's
improvement; and while they cannot be denied,
their proportionate representation in history are
sometimes dealt with as if they were the main
events of history, and were entitled to occupy the
largest space and foremost place in the picture
drawn by the historian. My purpose has been
something different from this; my desire has been
to make a picture of the marvellous changes which in
science and literature, by statesmen and philan-
thropists in the workshop of England during the
wonderful years of the nineteenth century. My
wish has been to make my readers acquainted
with the men who helped to bring about those
changes, as well as with the nature, extent, and
influence of the changes themselves; and thus to
tell the story of England's nineteenth century in
such a manner as to secure it an easy way of the
intellectual and a place in the history of the
my youngest readers.

Proper names – where in the book they appear



Fisheye view: one paragraph in full, and around it extracts

IN the volume and in the one to follow my purpose is to give an account of the social and political development of England since the opening of the century. Caricature is itself caricatured by the grotesque and ridiculous illustration which found such favour with Englishmen in the days of George III. the sketch which represents the war went on and on, until Wellington won the battle of Waterloo and then all was over.

Pitt's great opponent was Charles James Fox. Thus it was with England when the wished-for peace had at last been brought about. The common belief, not unnatural, was that with peace must come prosperity. Hardly any one now avows or admits himself to be a Tory, except as a sort of half-defiant joke. If a majority of the House of Commons decided against the policy of his favourite Minister, the King took not the slightest notice of the decision, but maintained it. He showed many times that he could act with promptness and decision in cases of sudden and unforeseen difficulty. The statesmen assembled at the Congress of Vienna concerned themselves just as little about the sentiments and the predilections of the populations with whom little more than half a century had passed before a Republic was again set up by the French people, and there does not now seem the slightest chance, come what may, of the influence they afterwards obtained over the Councils of reactionary tyrannies in France and Spain becoming the principal means of upsetting the whole fabric. Perhaps this was giving somewhat too wide an application to a principle sound and healthy within its limits; but it certainly is a principle which cannot be borne. But he knew enough to know that many of the evils of which Englishmen then complained were to be ascribed directly and almost altogether to a bad system of government. It is well to know that the injustice of Cochrane's conviction was recognized in the reign of William IV., by whom he was restored to his rank in the navy. The Prime Minister during a great part of Queen Victoria's time was Lord Liverpool, a man whose name will always be remembered as that of one of the most better opponents of the old system.

George IV. succeeded to the throne. The new King had ruled so long during the absence of his father that his formal elevation to the sovereign power did not make much change. George IV. had been brought up by his father on narrow, old-fashioned, stinted principles of education. He had a greater amount of natural ability than was given to George III., but he had not the elder King's purity of personal character. Something might have been made of George IV. under a better and more liberal sort of training in his early days; but the effort to oppress him or to turn him into a pattern slave proved, as under such conditions. His instincts and inclinations were generous; and he was at least capable of understanding a better political system than that which seemed perfect to the court. There must have been much charm of manner and some brilliancy in conversation and style of George IV., seeing that he soon became in his early days the closest friend of the Duke of Devonshire. It is out of all reason to suppose that such men as Fox and Sheridan could have clung to the companionship of a mere worthless profligate simply because he had a fine person. It is certain that at one time these men and others had great hopes that the accession of George IV. would prove a blessing to the cause of progress and to the nation.

The eyes of the Catholics turned to George IV. as to a man who had pledged himself in favour of their claims. The Irish people in general believed that he was likely to encourage some better system of government for Ireland than the mere rule of coercion laws and the strict administration of the law. There were, indeed, some Englishmen of advanced opinions who never trusted him from the first; but on the whole it may be taken for granted that there was much hope.

Of course the private life of George when Regent had been one of utter profligality and reckless dissipation. Now nothing can be more clear than this declaration, and the declaration, that the construction of the law, is what makes the Periton a great epoch in English history. A town or large village had received the right of sending representatives to the House of Commons.

Canning had laid down a policy which was absolutely new at the time but which became, with some slight and fitful deviations, the settled foreign policy of England. They might be good there, though he did not think they were, but they were bad elsewhere, because they were out of relation with the existing moral and social conditions. Yet Lord Eldon might have been convicted if he could only have known that although the Bill was destined to pass, it would, at all events, be made in itself an instrument of great constitutional crisis had arisen, and even Lord Eldon urged so far as to admit that the result of the Clare election must be to bring the Catholics into the House of Commons.

One of the Ordinances was that no newspaper should be offered for sale, or be allowed in any portion to leave the place in which it was printed, until five copies had been deposited with the Post Office. He had been complaining of the policy of the Duke of Wellington, and he suddenly said, looking fixedly at Peel, "Him I accuse not."

The writer of this book has often seen Moore's portrait in its place of honour in the house of the late Courtesan Fustell, Lord John's widow. We have to consider what must have been the effect of the robe delivery, the exquisitely chosen language, the appeals so fully addressed to ingrained prejudices. The Government, however, held firm; and the disfranchisement of what were called the rotten boroughs was absolutely accomplished so far as the House of Commons was concerned. So the King sent for Peel, but Peel saw that this was a case in which he could do nothing.

Most of them, even of those who actually went out and inspected their own estates, were willing to accept the reports of the managers and overseers whom they employed. The sermons of the day were often long and without any grown persons to stand upright, and the women had to creep on their hands and knees, crawling backward and forward. The attendance of flugging in the Navy is an event which comes well within the recollection of most of us.

Browsing and summarization

It was supposed to be easier, since we don't need to compose sentences.

We do make use of syntactic parsing (for phrase finding) and term weighting (to decide what's important)

However, summarization by picking sentences is difficult if the original document was not written with this in mind. Picking phrases seems to have the same problem.

Much of what we do is related to summarization

Can we test it?

Easy tasks to evaluate tend to be question-answering.
How could you tell whether someone has done well at "find a book that you will like"?

- a) User satisfaction; give them a questionnaire.
- b) User interest; give them a choice of browsing or solitaire.
- c) User competence; try to rate their success in some (manual) way.

What we're trying: give people a list of topics and ask them to rate a book as more or less valuable for different topics.

Conclusion

Large collections are coming (whether they will help quality is doubtful, but they will contain a lot of new information).

We need an interface oriented towards large books.

We still haven't explored the selection of materials in context.

We also need ways of judging bias, genre,...

And mostly we need to measure utility.